



25 April, 2024

Portfolio Piece: Build a Data Mart in SQL

Kirstin Moffatt (92128298)
DLBDSPBDM01




Table of Contents

<i>Abstract</i>	2
<i>Requirements Analysis</i>	2
What roles (person/user groups) are there?	2
What actions do these roles perform?	2
Which data and functions are required?	3
<i>Data Dictionary</i>	4
Profile	4
Address	4
Payment Details	5
Card Holder	5
Booking	6
User Verification Info	6
Wishlist	6
Calendar	7
Listing	7
Pricing.....	8
Photos	8
Reviews	8
Messaging.....	9
Amenities	9
Host Info.....	9
Guest Info.....	10
House Rules.....	10
User Profile Listing	10
Message Booking	11
Ratings Profile User	11
WishListing	11
Rules Availability	11
Amenities Availability	12
<i>Entity Relationship Diagram</i>	13

Abstract

This purpose of this project is to design a structured analytical data repository (OLAP) focusing on the infrastructure for the backend of the AirBnB interface. The goal is to ensure that the architecture integrates data from multiple operational systems providing a unified view of data assets enabling seamless data operations.

The first step involved collecting and understanding the business needs of AirBnB. Next was to translate the requirements into a technical solution starting with conceptualizing the data model. First, I defined two 'main' element tables from which I would then work outwards. The first element was *UserProfile* which would allow me to define different types of users of the system. The second element was *Listing* which would contain all the relevant data pertaining to making bookings. In structuring the relationship and table types, I took a holistic view of the project to ensure that the various components interact with each other in a seamless and logical manner.

Requirements Analysis

What roles (person/user groups) are there?

- Database administrator
- Data engineer
- Data architect and/or data analyst
- Business intelligence analyst and/or data analyst

What actions do these roles perform?

It is important to note that some of the actions performed by these roles can be transferred. Database modelling requires cross-functional coordination of roles and activities in order to achieve common objectives and goals. Data roles which have less direct involvement in database modelling, but where some responsibilities may overlap include a data analyst, business intelligence analyst and a project manager.

Database administrator: implements schema, grants access to database, monitors performance, ensures security and backups are handled ...

Data engineer: implements, develops and maintains database, writes SQL queries, optimizes performance, integrates database with other applications (in collaboration with software engineers) ...

Data architect: high-level design and strategic planning of the database, collaborates with stakeholders to translate business needs, ensures a smooth flow within the data table

Business Intelligence analyst: requirements gathering, define components of data mart, on-going support and integration of operational data within the system

Which data and functions are required?

The following table will help the user understand what function each elementary table holds.

Element	Fact table	Dimension table	Joining table	Aggregate Table
Profile		x		
Payment Details		x		
Card Holder		x		
Address		x		
User Verification Info	x			
Language Preferences			x	
Wish List		x		
Booking		x		
Calendar		x		
Listing		x		
Amenities Availability			x	
Amenities				
Photos				x
Review				x
Ratings Profile User	x			
Messaging				x
Host Info		x		
Guest Info		x		
Ratings Profile User			x	
User Profile Listing			x	
Messaging Booking			x	
Wish Listing			x	
Rule Availability			x	

Data Dictionary

Each table includes a unique identifier serving as the primary key for data retrieval and referencing purposes

Profile

Basic information pertaining to users on the platform including guests and hosts

	Column	Description	Data Type	NotNull	Unique
PK	UserID	A unique identifier for each user	int	x	x
	FirstName	The user's first name	varchar(200)	x	
	LastName	The user's last name	varchar(200)	x	
	UserName	A unique username as chosen by the signee	varchar(100)	x	x
	Email	The user's e-mail	varchar(100)	x	x
	Phone	The user's phone number	varchar(200)	x	x
	PhotoURL	The URL of the user's profile photo	varchar(200)	x	x
	PasswordHash	The hashed version of the user's password for secure storage	char(60)	x	x
FK	VerificationInfoID	Information related to the user's verification status	int	x	x
FK	RatingID	User generated reviews or feedback	int	x	x
FK	HostID	A unique identifier for user's who are hosts	int	x	x
FK	GuestID	A unique identifier for users who are guests	int	x	x
FK	AddressID	A unique identifier for each address	int	x	x
FK	TransactionID	A unique identifier for each payment transaction	int	x	x

Address

The address details of the user's

	Column	Description	Data Type	IsRequired	Unique
PK	AddressID	A unique identifier for each address	int	x	x
	Street	The street address, including house number and street name	varchar(200)	x	
	City	The city where the address is located	varchar(200)	x	
	StateProvince	The state or province where the address is located	varchar(200)	x	
	PostalCode	The postal or ZIP code for the address	int	x	
	Country	The country where the address is located	varchar(200)	x	

Payment Details

Information required to process transactions

	Column	Description	Data Type	IsRequired	Unique
PK	TransactionID	A unique identifier for each payment transaction	int	x	x
	PaymentAmount	The total amount of the payment	decimal	x	
	PaymentDate	The date the payment was made	date	x	
	CreditCardNumberHash	The credit card number used for the payment	char(20)	x	
	ExpirationDateHash	The expiration date of the credit card	char(10)	x	
	PaymentStatus	The current status of the payment (pending, completed, failed)	enum	x	
	CardVerificationCodeHash	The CVV (Card Verification Value) or security code of the credit card	char(3)	x	
FK	AddressID	The address associated with the credit card for billing purposes	int	x	
FK	CardholderID	The name of the credit card holder	Int	x	

Card Holder

Details of users who own credit or debit cards used for payments, facilitating secure and convenient transactions

	Column	Description	Data Type	IsRequired	Unique
PK	CardHolderID	A unique identifier for each cardholder	int	x	x
	Title	The title of the cardholder (Mr., Mrs., Miss.)	enum	x	
	FirstName	The cardholder's first name	varchar(200)	x	
	LastName	The cardholder's last name	varchar(200)	x	

Booking

A record of reservations made by users, containing necessary and additional booking-related information

	Column	Description	Data Type	IsRequired	Unique
PK	BookingID	A unique identifier for each booking	int	x	x
	CheckInDate	The date the user is scheduled to check in	date	x	
	CheckOutDate	The date the user is scheduled to check out	date	x	
	BookingStatus	The current status of the booking (pending, confirmed, cancelled)	enum	x	
FK	MessageID	Communication or messages related to the booking	int	x	x
FK	UserID	Reference to the user profile associated with the booking	int	x	x
FK	ListingID	A unique identifier for the listing being booked	int	x	x

User Verification Info

Tracking the verification status of users, indicating whether they have completed the required verification steps for authentication and trust purposes

	Column	Description	Data Type	IsRequired	Unique
PK	VerificationInfoID	A unique identifier for each set of verification information	int	x	x
	EmailVerified	Indicates whether the user's email address has been verified (true/false)	boolean	x	
	IDDocumentVerified	Indicates whether the user's identification document has been verified (true/false)	boolean	x	
	VerificationDate	The date when the verification was completed	date	x	

Wishlist

Collections of properties saved by users for future reference or consideration, allowing them to curate lists of desired accommodations

	Column	Description	Data Type	IsRequired	Unique
PK	WishlistID	A unique identifier for each wishlist	int	x	x
	CreationDate	The date the wishlist was created	date	x	
FK	ListingID	A unique ID associated with the listing	int	x	x
FK	GuestID	A unique identifier for the guest associated with creating the wishlist	int	x	

Calendar

Availability and booking information for properties on specific dates, enabling users to view and manage the property's availability and reservations

	Column	Description	Data Type	IsRequired	Unique
PK	CalendarID	A unique identifier for each calendar entry	int	x	x
FK	BookingID	A unique identifier for the booking associated with the calendar entry	int	x	

Listing

Details and attributes of properties available for booking on the platform

	Column	Description	Data Type	IsRequired	Unique
PK	ListingID	A unique identifier for each listing	int	x	x
	ListingName	The name or title of the listing	varchar(200)	x	x
	RoomType	The type of room or accommodation offered (apartment, private room, shared room, house)	enum	x	
	Description	A description of the listing, highlighting its features and attractions	text	x	
	MinimumStay	The minimum number of nights required for a stay at the listing	int	x	
	MaximumStay	The maximum number of nights allowed for a stay at the listing	int		
FK	AddressID	The address of the listing, which can include street, city, state/province, postal code, and country	int	x	x
FK	PricingID	Information related to pricing of the listing	int	x	
FK	PhotoID	URLs or references to photos showcasing the listing	int	x	x
FK	AmenityID	List of amenities available at the listing	int	x	
FK	RatingID	User-generated reviews related to the listing	int	x	
FK	Calendarid	Information related to the availability and booking calendar for the listing	int	x	
FK	HostID	A unique identifier linking the listing to the host	int	x	
FK	WishListID	A unique identifier linking the listing to a specific wishlist	int	x	x
FK	HouseRulesID	A unique identifier linking the listing to house rules	int	x	x

Pricing

Pricing details for properties including additional fee's and currency conversion

	Column	Description	Data Type	IsRequired	Unique
PK	PricingID	A unique identifier for each pricing entry	int	x	x
	Value	The numerical value of the pricing (amount per night)	numeric	x	
	Currency	The currency in which the pricing is specified	Varchar(20)	x	

Photos

Storing descriptions and metadata associated to photos of the listing

	Column	Description	Data Type	IsRequired	Unique
PK	PhotoID	A unique identifier for each photo	int	x	x
	PhotoData	Storing photo binary data	bytea	x	x
	Caption	A brief description or caption for the photo	varchar(200)		

Reviews

Feedback submitted by users, contributing to the reputation and trustworthiness of users

	Column	Description	Data Type	IsRequired	Unique
PK	RatingID	A unique identifier for each review rating	int	x	x
	Rating	The numeric rating given by the evaluator (e.g., on a scale of 1 to 5)	int	x	
	Comment	Optional textual comments or feedback provided by the evaluator	text	x	
FK	UserID	A unique identifier for the user who provided the review	int	x	

Messaging

Storing communication exchanges between users and hosts related to bookings, inquiries, and general interactions

	Column	Description	Data Type	IsRequired	Unique
PK	MessageID	A unique identifier for each message	int	x	x
	Content	The content or body of the message	varchar(300)	x	
	Timestamp	The date and time when the message was sent	timestamp	x	
FK	UserID	A unique identifier for the sender/ receiver of the message	int	x	
FK	BookingID	A unique identifier for the booking associated with the message	int	x	

Amenities

Features and facilities available in properties

	Column	Description	Data Type	IsRequired	Unique
PK	AmenityID	A unique identifier for each amenity	int	x	x
	AmenityName	The name or brief description of the amenity	varchar(200)	x	
	Category	The category or type to which the amenity belongs (basic, kitchen, connectivity, entertainment, safety&security, outdoor, additional)	enum	x	

Host Info

Information about hosts pertaining to status, response time and tenure on the platform

	Column	Description	Data Type	IsRequired	Unique
PK	HostID	A unique identifier for each host	int	x	x
	ResponseRate	An indication of the response rate of the host typically measured in percentages, representing how often a host will respond within 24 hours	numeric	x	
	SinceWhen	The date since when the host has been active or registered	date	x	
	HostStatus	The current status of the host (active, inactive, superhost)	enum	x	
FK	ListingID	A unique identifier linking the host to a listing	int	x	x

Guest Info

Information about guests pertaining to payment details, language preference, social media accounts and a customized wishlist

	Column	Description	Data Type	IsRequired	Unique
PK	GuestID	A unique identifier for each guest	int	x	x
	LanguagePreferences	The language preferences of the guest	enum	x	
	SocialMediaAccounts	Information about the guest's social media accounts (Google, Facebook, LinkedIn)	enum		
FK	AddressID	The address associated with the guest	int	x	
FK	TransactionID	Details related to the guest's payment methods or preferences	int	x	

House Rules

Rules pertaining to each listing

	Column	Description	Data Type	IsRequired	Unique
PK	HouseRulesID	A unique identifier for house rules	int	x	x
	RulesDescription	A brief description of each rule	text	x	

User Profile Listing

Relationship table linking user profiles to listings, facilitating efficient retrieval and analysis of review ratings associated with users

	Column	Description	Data Type	IsRequired	Unique
Composite Primary Key: ListingID, UserID					
FK	UserID	A unique identifier for the user	int	x	x
FK	ListingID	A unique identifier for the listing	int	x	x

Message Booking

Relationship table linking messages between users and bookings associated with listings

	Column	Description	Data Type	IsRequired	Unique
Composite Primary Key: BookingID, MessagingID					
FK	BookingID	A unique identifier for the booking associated with the messaging interaction	int	x	x
FK	MessagingID	A unique identifier for the messaging interaction related to the booking	int	x	x

Ratings Profile User

Relationship table linking user profiles to reviews, facilitating efficient retrieval and analysis of review ratings associated with users

	Column	Description	Data Type	IsRequired	Unique
Composite Primary Key: RatingsID, UserID					
FK	RatingID	A unique identifier for each rating	int	x	x
FK	UserID	A unique identifier for the user	int	x	x

WishListing

Relationship table linking wishlists to listings

	Column	Description	Data Type	IsRequired	Unique
Composite Primary Key: WishListID, ListingID					
FK	WishListID	A unique identifier for each wishlist	int	x	x
FK	ListingID	A unique identifier for the listing	int	x	x

Rules Availability

A list of rules available in the listing

	Column	Description	Data Type	IsRequired	Unique
Composite Primary Key: HouseRulesID, ListingID					
FK	HouseRulesID	A unique identifier for each rule	int	x	x

FK	ListingID	A unique identifier associated with the listing	int	x	x
	Available	Indicates whether the rule is active or not (true/false)	boolean	x	

Amenities Availability

A list of amenities available in the listing

	Column	Description	Data Type	IsRequired	Unique
Composite PK: AmenityID, ListingID					
FK	AmenityID	A unique identifier for each amenity	int	x	x
FK	ListingID	A unique identifier associated with the listing	int	x	x
	Available	Indicates whether the amenity is available or not (true/false)	boolean	x	

Entity Relationship Diagram

